

て信号が変動する場合や、平均的な特徴では信号を区別し難い場合等であっても、より高い精度で信号を検出できる高速信号探索方法について説明する。

【0049】図7は、本発明の第2実施形態による高速信号探索方法を適用した高速信号探索装置の構成例を示すブロック図である。本発明では、様々な処理対象信号を用いることができるが、ここでは、該処理対象信号の一例として、音響信号を用いる。本装置は、参照特徴量抽出手段201と入力特徴量抽出手段202と類似値計算手段203と移動量計算手段204とによって構成される。本装置は、予め登録された参照音響信号（即ち、見本となる検索したい音響信号）と入力音響信号（即ち、検索される音響信号）とを入力とし、参照音響信号と入力音響信号との類似値が所定の閾値 $\theta$ を越える箇所を入力音響信号から検出する。

【0050】ここで、図7に示される高速信号探索装置は、具体的には、CPU（中央処理装置）およびその周辺回路からなるコンピュータ装置で構成される。該コンピュータ装置は、所定の記録媒体（磁気ディスク、半導体メモリ等）に記録された制御プログラムで制御されることによって、図7に示される各手段として機能する。なお、上記制御プログラムは通信回線を介して頒布することが可能である。

【0051】図7において、参照特徴量抽出手段201は、参照音響信号から生成された特徴量系列上に注目窓（以下、「参照信号注目窓」と称する）を設定し、該参照信号注目窓を複数の注目窓（以下、「参照信号分割注目窓」と称する）に分割し、各参照信号分割注目窓内の特徴量系列を類似値計算手段203へ供給する。入力特徴量抽出手段202は、入力音響信号から生成された特徴量系列上に注目窓（以下、「入力信号注目窓」と称する）を設定し、該入力信号注目窓を複数の注目窓（以下、「入力信号分割注目窓」と称する）に分割し、各入力信号分割注目窓内の特徴量系列を類似値計算手段203へ供給する。

【0052】類似値計算手段203は、参照特徴量抽出手段201および入力特徴量抽出手段202から供給された特徴量系列に基づいて、各参照信号分割注目窓内の特徴量系列と（該参照信号分割注目窓に対応する）各入力信号分割注目窓内の特徴量系列との類似値（以下、「入力信号類似値」と称する）を計算する。そして、類似値計算手段203は、計算された入力信号類似値が閾値 $\theta$ より大きいかな否かに基づいて、入力信号注目窓が現在示す位置に参照音響信号が存在するか否かを判定する。入力信号注目窓が現在示す位置に参照音響信号が存在する場合、類似値計算手段203は、入力信号注目窓が現在示す位置（時刻等）を信号検出結果として出力する。又、類似値計算手段203は、入力信号注目窓の移動可能量を計算するために必要な類似値を、移動量計算手段204へ供給する。

【0053】移動量計算手段204は、類似値計算手段203から供給された類似値に基づいて、該類似値に対応する入力信号注目窓の近傍の入力信号注目窓に関する類似値の上限値を計算する。そして、移動量計算手段204は、その上限値に基づいて、入力信号注目窓の移動可能量を計算し、該移動可能量を入力特徴量抽出手段202へ供給する。これによって、入力特徴量抽出手段202は、入力信号注目窓の位置を、移動量計算手段204から供給された移動可能量だけ移動させる。このように、本装置は、入力信号注目窓を順次移動させながら、移動後の入力信号注目窓内の入力音響信号から生成される特徴量系列について、上記各処理を同様に繰り返す。

【0054】次に、図8および図9を参照して、本装置の動作を具体的に説明する。参照特徴量抽出手段201は、初めに、予め登録された参照音響信号を読み込む。次に、参照特徴量抽出手段201は、読み込まれた参照音響信号について、特徴量の抽出を行う（ステップS201）。本発明では、様々な特徴量を用いることができるが、ここでは、該特徴量の一例として、スペクトル特徴を用いる。この場合、特徴量の抽出は、例えば、帯域通過フィルタを用いて行うことができる。

【0055】特徴量抽出の具体例を以下に説明する。例えば、テレビやラジオ等の放送信号から10秒程度の特定の音響信号を検索したい場合、図10に示す装置で特徴量を抽出すると、良い結果が得られる。即ち、7個の帯域通過フィルタを該帯域通過フィルタの中心周波数が対数軸上で等間隔になるように設定し、各帯域通過フィルタの出力波形に対して60ミリ秒の時間窓を設定し、その時間窓内における出力波形の自乗の平均値を計算する。このようにして得られた7個の平均値を一組にして7次元特徴ベクトルとする。この時間窓を12ミリ秒ずつずらしながら、順次特徴ベクトルの算出を行う。この場合、特徴ベクトルは、12ミリ秒毎に1つずつ得られる。参照特徴量抽出手段201においては、このようにして、参照音響信号の各周波数帯域の成分を要素とする特徴ベクトルが、時系列的に順次得られる。なお、上記特徴量抽出方法としては、この他に、以下の方法も考えられる。その方法とは、即ち、7個の帯域通過フィルタを該帯域通過フィルタの中心周波数が対数軸上で等間隔になるように設定し、各帯域通過フィルタの出力を12ミリ秒毎に60ミリ秒幅で取り出し、取り出された出力とその直前に取り出された出力との差分を計算し、該差分の自乗の平均値を計算し、得られた7個の平均値を一組にして7次元特徴ベクトルとする方法である。

【0056】続いて、参照特徴量抽出手段201は、それら特徴ベクトルの時系列に対して注目窓を設定する（ステップS202）。参照特徴量抽出手段201は、まず、参照音響信号全体に対して1つの注目窓（参照信号注目窓）を設定する（図8中、参照特徴量抽出手段201内の「注目窓」参照）。尚、図8には、「参照信号

31

注目窓内の横方向が時間軸に相当し、該時間軸方向において特徴ベクトルが順次得られる」ことが模式的に示されている。

【0057】次に、参照特徴量抽出手段201は、参照信号注目窓を、時間軸方向において、複数の参照信号分割注目窓に分割する（ステップS203、図8の参照特徴量抽出手段201内の“分割”参照）。尚、この分割数は、処理対象とする信号の形態等に応じて適宜定められ得るが、ここでは、例えば8分割とする。これにより、参照特徴量抽出手段201は、各参照信号分割注目窓に含まれる特徴量（即ち、特徴ベクトルの時系列）を、順次、類似値計算手段203へ供給する。

【0058】入力特徴量抽出手段202は、初めに、入力音響信号を読み込む。次に、入力特徴量抽出手段202は、読み込まれた入力音響信号について、特徴量の抽出を行う（ステップS204）。ここで、入力特徴量抽出手段202は、参照特徴量抽出手段201と同じ方法（図10参照）で、特徴量（即ち、7次元特徴ベクトルの時系列）を抽出する。故に、ここでは、その説明を省略する。

【0059】続いて、入力特徴量抽出手段202は、それら特徴ベクトルの時系列に対して注目窓を設定する（ステップS205）。入力特徴量抽出手段202は、まず、上記参照信号注目窓と同じ長さの注目窓（入力信号注目窓）を入力音響信号に対して設定する（図8中、入力特徴量抽出手段202内の“注目窓”参照）。尚、図8には、「入力信号注目窓内の横方向が時間軸に相当し、該時間軸方向において特徴ベクトルが順次得られる」ことが模式的に示されている。

【0060】尚、入力信号注目窓の位置は、処理開始当初は、入力音響信号から作成された特徴量系列の先頭に設定されるが、処理の進行と共に、後述する方法によって、該特徴量系列上を時間軸方向に順次移動される。この移動量は、移動量計算手段204によって計算される。

【0061】続いて、入力特徴量抽出手段202は、入力信号注目窓を、時間軸方向において、複数の入力信号分割注目窓に分割する（ステップS206、図8の入力特徴量抽出手段202内の“分割”参照）。尚、入力信号注目窓の分割数は、参照信号注目窓の分割数と等しい。これにより、入力特徴量抽出手段202は、各入力信号分割注目窓に含まれる特徴量（即ち、特徴ベクトルの時系列）を、順次、類似値計算手段203へ出力する。

【0062】類似値計算手段203は、参照特徴量抽出手段201から、参照信号分割注目窓で分割された特徴ベクトルの時系列を順次読み、かつ、入力特徴量抽出手段202から、入力信号分割注目窓で分割された特徴ベクトルの時系列を順次読み込む。

【0063】次に、類似値計算手段203は、参照信号

32

分割注目窓内の特徴ベクトルの時系列に基づいて、該特徴ベクトルのヒストグラムを作成し（ステップS207）、かつ、入力信号分割注目窓内の特徴ベクトルの時系列に基づいて、該特徴ベクトルのヒストグラムを作成する（ステップS208）。これらのヒストグラムは、いずれも、特徴ベクトルの各要素が取りうる値の範囲（最小値から最大値まで）を複数のビン（区間）に分割することによって作成される。例えば、各要素が取りうる値の範囲を3つの区間（仮に、「区間A、B、C」とする）に分割し、かつ、各特徴ベクトルの要素数が7であるとすれば、7個の要素のそれぞれは区間A、B、Cのうちのいずれかに属することになる。従って、1つの特徴ベクトルについて考えると、該特徴ベクトルの要素の組み合わせとしては、7個の要素が全て区間Aに属することを示す（A、A、A、A、A、A、A）から、7個の要素が全て区間Cに属することを示す（C、C、C、C、C、C、C）まで、合計で3の7乗通りの組み合わせが考えられる。以上のことから、特徴ベクトルのヒストグラムの全体のビン数（即ち、ヒストグラムの横軸に配置される区間の数）は3の7乗個となる。従って、このようにヒストグラムの横軸を設定する場合には、各特徴ベクトルは、この3の7乗個の区間のうちのどれか一つに分類されることとなる。

【0064】以上述べた方法によって、類似値計算手段203は、参照特徴量抽出手段201から供給された（時間軸方向に分割された）特徴ベクトルと入力特徴量抽出手段202から供給された（時間軸方向に分割された）特徴ベクトルとについて、それぞれヒストグラムを作成する。ここでは、これらのヒストグラムを、 $G^1$ 、 $G^2$ 、 $\dots$ 、 $G^n$ および $H^1$ 、 $H^2$ 、 $\dots$ 、 $H^n$ とする。ただし、 $n$ は元の注目窓（参照信号注目窓および入力信号注目窓）の分割数であり（以下、1～ $n$ を「分割番号」と称する）、 $G$ は参照音響信号の特徴ベクトルから作成されたヒストグラム（以下、「参照信号ヒストグラム」と称する）を示し、 $H$ は入力音響信号の特徴ベクトルから作成されたヒストグラム（以下、「入力信号ヒストグラム」と称する）を示す。

【0065】続いて、類似値計算手段203は、分割番号が等しい参照信号ヒストグラムと入力信号ヒストグラムとの類似値を計算する（ステップS209）。本発明では、様々な類似値を用いることができるが、ここでは、該類似値の一例として、重なり類似値を用いる。ここで、分割番号 $k$ における（参照信号ヒストグラムと入力信号ヒストグラムとの）類似値 $S_k$ は、以下の式によって定義される。

【0066】

【数3】

$$S_k = \frac{1}{D_k} \sum_{j=1}^L \min(g_j^k, h_j^k)$$

50

【0067】ここで、 $D_k$ は分割番号 $k$ のヒストグラムの総度数であり、 $L$ はヒストグラムの区間数（上記の例では3の7乗）であり、 $g^k_j$ は $k$ 番目のヒストグラム $G^k$ の $j$ 番目の区間の値であり、 $h^k_j$ は $k$ 番目のヒストグラム $H^k$ の $j$ 番目の区間の値であり、 $\min(g^k_j, h^k_j)$ は $g^k_j$ と $h^k_j$ とにおける小さい方の値を示す。

【0068】また、注目窓全体における類似値（即ち、参照信号注目窓と入力信号注目窓とにおける類似値） $S$ は、以下の式によって定義される（以下、該類似値 $S$ を「全体類似値」と称する）。

【0069】

【数4】

$$S = \min(S_1, S_2, \dots, S_n)$$

【0070】ここで、 $\min(S_1, S_2, \dots, S_n)$ は、 $S_1, S_2, \dots, S_n$ のうちの最小値を表す。

【0071】類似値 $S_k$ の計算は、1つずつ（例えば、値 $k$ の小さい順に）行われる。もし、計算された類似値 $S_k$ が閾値 $\theta$ 以下であれば、数4の式に基づいて、全体類似値 $S$ の最小値が閾値 $\theta$ 以下であることは明らかであるので、後続の類似値 $S_{k+1}, S_{k+2}, S_{k+3}, \dots$ の計算はもはや行われる必要がない。そして、類似値計算手段203は、それまでに計算された類似値のうちの最小値を、移動量計算手段204に供給する。

【0072】一方、もし、すべての類似値 $S_1 \sim S_n$ が閾値 $\theta$ より大であれば、全体類似値 $S$ は閾値 $\theta$ より大となる。このことは、「入力音響信号（の特徴ベクトル）に対する入力信号注目窓の現在位置に、参照音響信号が検出された」ということを意味する。そこで、類似値計算手段203は、この現在位置（時刻）を、信号検出結果の一部として出力する。また、この場合も、類似値計算手段203は、全体類似値 $S$ を、移動量計算手段204に供給する。

【0073】移動量計算手段204は、はじめに、類似値計算手段203から供給された全体類似値 $S$ を読み込む。次に、移動量計算手段204は、移動可能量 $w$ を計算する（ステップS210）。ここで、移動可能量 $w$ は、以下の式によって求められる。

【0074】

【数5】

$$w = \begin{cases} \text{floor}(D(\theta - S)) + 1 & (S < \theta \text{ のとき}) \\ 1 & (\text{上記以外}) \end{cases}$$

【0075】ここで、移動可能量 $w$ の単位は、特徴ベクトルの個数であり、 $\text{floor}()$ は、小数点以下の切り下げを表し、 $D$ は、 $S = S_k$ を満たす $k$ 番目の分割注目窓（参照信号分割注目窓および入力信号分割注目窓）内の特徴ベクトルのヒストグラムの総度数であり、 $\theta$ は上記閾値である。

【0076】数5の式は、「現時点で $S < \theta$ ならば、特徴ベクトル $(w-1)$ 個分だけ入力信号注目窓を移動させても、全体類似値 $S$ は決して閾値 $\theta$ を越えない」ということを意味している。これは、入力信号注目窓を移動させたときに、入力信号注目窓の外に出ていく特徴ベクトルが全てヒストグラムの重なりに寄与していないものであり、かつ、入力信号注目窓の中に入ってくる特徴ベクトルが全てヒストグラムの重なりに寄与するものである場合（即ち、全体類似値 $S$ が最も早く閾値 $\theta$ に達する場合）」を考えれば、容易に理解される。即ち、このような場合に類似値が最も増加するので、このような場合を想定すると、入力信号注目窓を特徴ベクトル $(w-1)$ 個分だけ移動させたときの全体類似値 $S$ の上限値は閾値 $\theta$ となる。このことから、 $S < \theta$ ならば、全体類似値 $S$ が閾値 $\theta$ を越える可能性がでてくる特徴ベクトル $w$ 個分の移動量を移動可能量とするのである。

【0077】一方、 $S \geq \theta$ ならば、類似値のローカルピークを見い出すために、 $w = 1$ とし、注目窓のスキップを行うことなく検索を行う。

【0078】移動量計算手段204から出力された移動可能量 $w$ は、入力特徴量抽出手段202に供給される。入力特徴量抽出手段202は、入力信号注目窓を、特徴ベクトル $w$ 個分だけ移動させる（ステップS211、図8の入力特徴量抽出手段202内の破線の注目窓参照）。以後、本装置は、上述した処理（複数の分割注目窓の設定、ヒストグラムの作成、類似値の計算、移動可能量の計算等）を同様に繰り返す（ステップS212）。また、入力信号注目窓の移動によって、入力信号注目窓が入力音響信号（の特徴ベクトルの時系列）の終端を越えたならば、全入力音響信号を探索したことになるので、探索処理を終了する。

【0079】次に、本実施形態による高速信号探索装置の動作実験例を示す。本装置をワークステーション（SGI社O<sub>2</sub>）に実装し、テレビ放送を題材として音響信号の検出を行った結果を図11と図12に示す。いずれも、縦軸は類似値、横軸は時刻を示す。これらは、あるコマーシャル（15秒間）を参照信号とし、実際のテレビ放送の録音（1998年1月22日、18時22分から6時間分）を入力信号として検出を行ったものである。なお本実験では、サンプリング周波数=11.025kHz、特徴ベクトルの次元=7、各特徴ベクトルの要素のピン数=3とした。

【0080】図11は、注目窓を分割しなかった場合、すなわち本方法を実行しなかった場合、図12は、本方法を実行して注目窓を8分割した場合である。それぞれの図において、○印を付けた部分が本装置によって検索された箇所を示す。人手によって確かめたところによれば、本実験における正解は、18:48、23:22、00:11の3箇所であった。図11では、正解箇所の他に、3箇所余分に検出しているのに対し、図12で

は、正しい検出結果が得られている。

【0081】また、一般に、検索すべき箇所における類似値と、検索されるべきでない箇所における類似値の比が大きいほど、設定値の設定に対するマージンが大きくなるため、安定した検索が可能である。本実験では、図12の場合の方が、図11の場合に比べて前記比の値が大きくなっており、注目窓の分割による効果が表れていることが分かる。

【0082】また、画像パターンに対して信号検出を行った結果を図13と図14に示す。いずれも、縦軸は類似値、横軸は時刻を示す。これらは、図11および図12とは別のコマーシャル（15秒間）の画像から得た色情報を参照信号とし、実際のテレビ放送を録画して得た色情報（1998年1月22日、18時22分から6時間分）を入力信号として検出を行ったものである。なお本実験では、特徴ベクトルの次元=3、各特徴ベクトルの要素のピン数=8とした。

【0083】図13は、注目窓を分割しなかった場合（即ち、本方法を実行しなかった場合）、図14は、本方法を実行して注目窓を8分割した場合である。それぞれにおいて、○印を付けた部分が本装置によって検索された箇所を示す。人手によって確かめたところによれば、本実験における正解は、18:48、19:56、20:35、22:54の4箇所であった。図13では、正解箇所のうちの一つ（20:35）が検出できておらず、また、2箇所余分に検出しているのに対し、図14では、正しい検出結果が得られている。

#### 【0084】§3. 第3実施形態

次に、本発明の第3実施形態について図面を参照して説明する。

【0085】第1実施形態および第2実施形態では、予め登録された参照信号に類似した信号の位置を検出する方法が説明された。しかしながら、これらの方法には、「多数の参照信号に基づいて信号を検出する場合、参照信号の数だけ検出処理を繰り返す必要があり、計算量が増大して、検出速度が遅くなる」という課題がある。

【0086】そこで、第3実施形態では、多数の参照信号に基づいて信号を検出する場合であっても、従来より少ない処理で信号を検出できる高速信号探索方法について説明する。

【0087】図15は、本発明の第3実施形態による高速信号探索方法を適用した高速信号探索装置の構成例を示すブロック図である。本発明では、様々な処理対象信号を用いることができるが、ここでは、該処理対象信号の一例として、音響信号を用いる。図15において、本装置1は、参照特徴量計算手段310と入力特徴量計算手段312と参照信号間類似値計算手段314と類似値計算手段316と移動可能量計算手段318とによって構成される。

【0088】ここで、図15に示される高速信号探索装

置は、具体的には、CPU（中央処理装置）およびその周辺回路からなるコンピュータ装置で構成される。該コンピュータ装置は、所定の記録媒体（磁気ディスク、半導体メモリ等）に記録された制御プログラムで制御されることによって、図15に示される各手段として機能する。なお、上記制御プログラムは通信回線を介して頒布することが可能である。

【0089】本装置1は、予め登録された参照音響信号（即ち、見本となる検索したい音響信号）と入力音響信号（即ち、検索される音響信号）とを入力とし、参照音響信号と入力音響信号との類似値が所定の閾値 $\theta$ を越える箇所を入力音響信号から検出する。

【0090】参照特徴量計算手段310は、複数の参照音響信号について、特徴量系列を生成する。また、入力特徴量計算手段312は、入力音響信号から特徴量系列を生成し、該特徴量系列上に注目窓（以下、「入力信号注目窓」と称する）を設定する。

【0091】参照信号間類似値計算手段314は、各参照音響信号（の特徴量系列）同士の類似値（以下、「参照信号間類似値」と称する）を計算する。類似値計算手段316は、参照特徴量計算手段310で生成された特徴量系列と入力特徴量計算手段312で設定された入力信号注目窓内の特徴量系列との類似値（以下、「入力信号類似値」と称する）を計算する。

【0092】移動可能量計算手段318は、参照信号間類似値計算手段314で計算された参照信号間類似値と類似値計算手段316で計算された入力信号類似値とに基づいて、該入力信号類似値に対応する入力信号注目窓の近傍の入力信号注目窓における（入力音響信号と参照音響信号との間の）入力信号類似値の範囲を規定するパラメータ（例えば、類似値の上限値）を計算する。そして、移動可能量計算手段318は、該パラメータの値に基づいて、入力信号注目窓の移動可能量を計算する。

【0093】次に、図16を参照して、本装置の動作を具体的に説明する。説明を簡明にするために、以下の説明では、全ての参照音響信号の長さは等しいものとする。なお実際には、本発明は、必ずしも全ての参照音響信号の長さが等しくない場合においても適用できる。

【0094】参照特徴量計算手段310は、初めに、与えられた全ての参照音響信号を読み込む。次に、参照特徴量計算手段310は、読み込まれた参照音響信号について、特徴量の抽出を行う（ステップS301）。本発明では、様々な特徴量を用いることができるが、ここでは、該特徴量の一例として、スペクトル特徴を用いる。この場合、参照特徴量計算手段310は、第2実施形態の参照特徴量抽出手段201と同じ方法（図10参照）で、特徴量（即ち、7次元特徴ベクトルの時系列）を抽出する。故に、ここでは、その説明を省略する。

【0095】次に、参照特徴量計算手段310は、上記特徴ベクトルの時系列に基づいて、該特徴ベクトルのヒ

10

20

30

40

50

37

ストグラムを作成する（ステップS302）。ここで、該ヒストグラムは、特徴ベクトルの各要素が取りうる値の範囲（最小値から最大値まで）を複数のビン（区間）に分割することによって作成される。参照特徴量計算手段310は、類似値計算手段203（第2実施形態）と同様の方法で、ヒストグラムを作成する。故に、ここでは、その説明を省略する。

【0096】なお、本実施形態では、1つの参照音響信号（の特徴ベクトルの時系列）の全体についてヒストグラムを作成しているが、この他にも、第2実施形態に示されるように、参照音響信号（の特徴ベクトルの時系列）を時間方向に分割し、分割後の各参照音響信号（の特徴ベクトルの時系列）についてヒストグラムを作成してもよい。分割の数は、例えば4とすると良い。この場合は、一つの参照音響信号について4つのヒストグラムが作成されることになる。

【0097】入力特徴量計算手段312は、初めに、入力音響信号を読み込む。次に、入力特徴量計算手段312は、読み込まれた入力音響信号について、特徴量の抽出を行う（ステップS303）。ここで、入力特徴量計算手段312は、参照特徴量計算手段310と同様の方法で、特徴量を抽出する。

【0098】続いて、入力特徴量計算手段312は、参照特徴量計算手段310に与えられた参照音響信号と同じ長さの入力信号注目窓を、抽出された特徴量（特徴ベクトルの時系列）に対して設定する（ステップS304）。尚、入力信号注目窓の位置は、処理開始当初は、入力音響信号から作成された特徴量系列の先頭に設定されるが、処理の進行と共に、後述する方法によって、該特徴量系列上を時間軸方向に順次移動される。この移動可能量は、類似値計算手段316によって計算される。

【0099】次に、入力特徴量計算手段312は、入力信号注目窓内の特徴ベクトルの時系列に基づいて、該特徴ベクトルのヒストグラムを作成する（ステップS305）。ここで、該ヒストグラムは、特徴ベクトルの各要素が取りうる値の範囲（最小値から最大値まで）を複数のビン（区間）に分割することによって作成される。入力特徴量計算手段312は、参照特徴量計算手段310と同様の方法で、ヒストグラムを作成する。故に、ここでは、その説明を省略する。

【0100】なお、参照特徴量計算手段310が（第2実施形態と同様に）参照音響信号を時間軸方向に分割した場合、入力特徴量計算手段312も、同様に、入力信号注目窓（内の特徴ベクトルの時系列）を時間軸方向に分割する。この分割数は、参照特徴量計算手段310における分割数と同じとする。従って、参照特徴量計算手段310において1つの参照音響信号から作成されたのと同数のヒストグラムが、入力信号注目窓内の入力音響信号から作成される。

【0101】参照信号間類似値計算手段314は、初め

38

に、参照特徴量計算手段310から供給されるヒストグラム（各参照音響信号のヒストグラム）を読み込む。ここでは、説明を簡明にするため、参照音響信号が分割されていない場合を考える（ただし、これは添字の扱いが複雑になることを防ぐためであり、本発明は、参照音響信号が分割される場合に対しても適用可能である）。ここでは、参照特徴量計算手段310から供給されるヒストグラムを、 $G^1, G^2, \dots, G^N$ とする。但し、 $N$ は参照音響信号の個数である。

【0102】参照信号間類似値計算手段314は、 $N$ 個の参照音響信号から2つの参照音響信号を取り出す全ての組み合わせについて、取り出された2つの参照音響信号間の類似値を計算する（ステップS306）。ここで、 $k$ 番目の参照音響信号のヒストグラム $g$ と $m$ 番目の参照音響信号のヒストグラム $h$ との類似値 $S(g, h)$ は、以下の式によって定義される。

【0103】

【数6】

$$S(g, h) = \frac{1}{D} \sum_{j=1}^L \min(g_j, h_j)$$

【0104】ここで、 $D$ はヒストグラムの総度数であり、 $L$ はヒストグラムの区間数（上記の例では3の7乗）であり、 $g_j$ はヒストグラム $g$ の $j$ 番目の区間の値であり、 $h_j$ はヒストグラム $h$ の $j$ 番目の区間の値であり、 $\min(g_j, h_j)$ は $g_j$ と $h_j$ とにおける小さい方の値を示す。

【0105】なお、参照特徴量計算手段310が（第2実施形態と同様に）参照音響信号を時間軸方向に分割した場合、参照信号間類似値計算手段314は、2つの参照音響信号の互に対応する分割部分について類似値計算を行う。該類似値計算の結果は、記憶手段（図示略）に格納され、移動可能量計算手段318に供給される。

【0106】類似値計算手段316は、初めに、参照特徴量計算手段310から供給されるヒストグラムと入力特徴量計算手段312から供給されるヒストグラムとを読み込む。ここでは、説明を簡明にするため、参照音響信号および入力音響信号が分割されていない場合を考える（ただし、これは添字の扱いが複雑になることを防ぐためであり、本発明は、参照音響信号および入力音響信号が分割される場合に対しても適用可能である）。ここでは、参照特徴量計算手段310から供給されるヒストグラムを、 $G^1, G^2, \dots, G^N$ とし、入力特徴量計算手段312から供給されるヒストグラムを $H$ とする。但し、 $N$ は参照音響信号の個数である。

【0107】続いて、類似値計算手段316は、参照音響信号のヒストグラムを1つ選択し、選択されたヒストグラムと入力音響信号のヒストグラムとの類似値を計算する（ステップS307）。参照音響信号のヒストグラム $g$ と入力音響信号のヒストグラム $h$ との類似値 $S$

10

20

30

40

50



39

(g, h) は、以下の式によって定義される。

【0108】

【数7】

$$S(g, h) = \frac{1}{D} \sum_{j=1}^L \min(g_j, h_j)$$

【0109】ここで、Dはヒストグラムの総度数であり、Lはヒストグラムの区間数（上記の例では3の7乗）であり、 $g_j$ はヒストグラムgのj番目の区間の値であり、 $h_j$ はヒストグラムhのj番目の区間の値であり、 $\min(g_j, h_j)$ は $g_j$ と $h_j$ における小さい方の値を示す。

【0110】本実施形態では、N個の参照音響信号が入力されるが、類似値の計算は、一度に、参照音響信号1つだけに対して行われる。類似値の計算に使用される参照音響信号の選択にあたっては、その時点における類似値の上限値が閾値 $\theta$ を越える可能性がある参照音響信号が選択される。どの参照音響信号がこれに該当するかは、後に説明する移動可能量計算手段318において判明するので、類似値計算手段316は、移動可能量計算手段318から、その情報を取得する。

【0111】なお、（第2実施形態と同様に）参照音響信号および入力音響信号が時間軸方向に分割されている場合には、類似値の計算は、（分割された）参照音響信号および入力音響信号の1つの分割部分に対してだけ行われる。

【0112】類似値計算手段316は、計算された類似値を、移動可能量計算手段318に供給する。これに加えて、もし、類似値が閾値 $\theta$ を越えている場合（参照音響信号および入力音響信号が時間軸方向に分割されている場合には、全ての分割分について、類似値が閾値 $\theta$ を越えている場合）、このことは、「その参照音響信号が入力音響信号中に存在している」ことを意味するので、類似値計算手段316は、信号検出結果として、参照音響信号の番号と入力信号注目窓の現在位置とを出力する。

【0113】移動可能量計算手段318は、初めに、入力音響信号とk番目の参照音響信号との類似値 $S^k$ を、類似値計算手段316から読み込む。次に、移動可能量計算手段318は、移動可能量 $w_k$ を計算する（ステップS308）。ここで、移動可能量 $w_k$ は、以下の式によって求められる。

【0114】

【数8】

$$w_k = \begin{cases} \text{floor}(D(\theta - S^k)) + 1 & (S^k < \theta \text{ のとき}) \\ 1 & (\text{上記以外}) \end{cases}$$

【0115】ここで、移動可能量 $w_k$ の単位は、特徴ベクトルの個数であり、 $\text{floor}()$ は、小数点以下の切り

40

下げを表し、Dは、 $S = S^k$ を満たすk番目の参照音響信号の特徴ベクトルのヒストグラムの総度数であり、 $\theta$ は上記閾値である。

【0116】数8の式は、「現時点において $S^k < \theta$ ならば、特徴ベクトル( $w_k - 1$ )個分だけ入力信号注目窓を移動させても、類似値 $S^k$ は決して閾値 $\theta$ を越えない」ということを意味している。これは、「入力信号注目窓を移動させたときに、入力信号注目窓の外に出ていく特徴ベクトルが全てヒストグラムの重なりに寄与していないものであり、かつ、入力信号注目窓の中に入ってくる特徴ベクトルが全てヒストグラムの重なりに寄与するものである場合（即ち、類似値 $S^k$ が最も早く閾値 $\theta$ に達する場合）」を考えれば、容易に理解される。即ち、このような場合に類似値が最も増加するので、このような場合を想定すると、入力信号注目窓を特徴ベクトル( $w_k - 1$ )個分だけ移動させたときの類似値 $S^k$ の上限値は閾値 $\theta$ となる。

【0117】一方、 $S^k \geq \theta$ ならば、類似値のローカルピークを見出すために、 $w_k = 1$ とし、注目窓のスキップを行うことなく検索を行う。

【0118】上述した処理は、1つの（即ち、k番目の）参照音響信号に対する処理である。参照音響信号が多数存在する場合、従来の方法では、この処理を参照音響信号の数だけ単純に繰り返すことによって、各参照音響信号に関する移動可能量を求める（これを反復法と呼ぶ）。これに対して、本発明は、k番目の参照音響信号に関する類似値 $S^k$ が得られると、該類似値 $S^k$ に基づいて、k番目以外の参照音響信号に関する移動可能量も求めることを特徴とする。故に、本発明によれば、上記反復法よりも類似値計算回数を削減することができる。

【0119】類似値計算回数を削減する方法について以下に説明する。2つの参照音響信号 $R^1$ 、 $R^2$ からそれぞれヒストグラム $G^1$ 、 $G^2$ が作成され、入力音響信号IからヒストグラムHが作成されているとする。但し、ヒストグラム $G^1$ 、 $G^2$ の総度数は等しく、かつ、参照音響信号および入力音響信号は時間軸方向に分割されていないものとする。また、ヒストグラム $G^1$ とヒストグラムHとの入力信号類似値 $S^1$ は既に得られているとする。

【0120】このとき、もし、ヒストグラム $G^1$ とヒストグラム $G^2$ とがよく似ているならば、ヒストグラムHとヒストグラム $G^2$ との入力信号類似値 $S^2$ を計算することなく、入力信号類似値 $S^1$ の上限値が得られる。即ち、ヒストグラム $G^1$ とヒストグラム $G^2$ との参照信号間類似値を $S^{12}$ とすると、ヒストグラム $G^1$ とヒストグラム $G^2$ との全ての不一致部分が入力信号類似値 $S^1$ の増加に寄与する場合に、入力信号類似値 $S^2$ はその上限値に達する。このことは、以下の式によって表される。

【0121】

【数9】

50

41

$$S^m \leq S^k + (1 - S^k)$$

【0122】但し、元々 $S^* \leq 1$ であるので、数9の式は、 $S^* < S^{**}$ のときのみ有効である。

【0123】一方、もし、ヒストグラム $G^*$ とヒストグラム $H$ とがよく似ているならば、ヒストグラム $G^*$ とヒストグラム $H$ との全ての不一致部分が入力信号類似値 $S^m$ の増加に寄与する場合に、入力信号類似値 $S^*$ はその上限値に達する。このことは、以下の式によって表される。

【0124】

【数10】

$$S^m \leq S^{km} + (1 - S^k)$$

【0125】但し、元々 $S^* \leq 1$ であるので、数10の式は、 $S^* > S^{**}$ のときのみ有効である。数9の式と数10の式とをまとめると、次式が得られる。

【0126】

【数11】

$$S^m \leq 1 - |S^k - S^{km}|$$

【0127】即ち、入力信号類似値 $S^*$ と参照信号間類似値 $S^{**}$ との差が大きければ大きいほど、入力信号類似値 $S^*$ の上限値は小さくなる。このとき、移動可能量 $w$ は、以下の式によって計算される。

【0128】

【数12】

$$w_m = \begin{cases} \text{floor}(D(\theta - S^m)) + 1 & (S^m < \theta \text{ のとき}) \\ 1 & (\text{上記以外}) \end{cases}$$

【0129】即ち、入力信号類似値 $S^*$ の上限値が小さければ小さいほど、移動可能量 $w$ は大きくなる。

【0130】以上述べたように、本発明によれば、2つの参照音響信号 $R^*$ 、 $R^*$ の参照信号間類似値 $S^{**}$ が既知である場合、参照音響信号 $R^*$ と入力音響信号 $I$ との入力信号類似値 $S^*$ が得られると、該入力信号類似値 $S^*$ に基づいて、参照音響信号 $R^*$ に関する移動可能量だけでなく、参照音響信号 $R^*$ に関する移動可能量をも求めることができる。そこで、 $k$ 番目の参照音響信号 $R^*$ と入力音響信号 $I$ との入力信号類似値 $S^*$ が（類似値計算手段316において）得られる度に、移動可能量計算手段318は、参照音響信号 $R^*$ と該参照音響信号 $R^*$ 以外の参照音響信号との参照信号間類似値を（参照信号間類似値計算手段314から）取り出し、該参照信号間類似値および数11、数12の式を用いて、（入力信号注目窓の）移動可能量を、全ての参照音響信号に関して求める。そして、このようにして求められた移動可能量がそれ以前に求められた移動可能量よりも大きい場合、移動可能量計算手段318は、移動可能量を更新する。

42

【0131】類似値計算手段316は、更新された移動可能量 $w_i$ （ $i = 1, \dots, N$ ）を、入力特徴量計算手段312に供給する。入力特徴量計算手段312は、移動可能量 $w_i$ のうちの最も小さい移動可能量だけ、入力信号注目窓を移動させる（ステップS309）。以後、本装置は、上述した処理（ヒストグラムの作成、類似値の計算、移動可能量の計算等）を同様に繰り返す（ステップS310）。また、入力信号注目窓の移動によって、入力信号注目窓が入力音響信号（の特徴ベクトルの時系列）の終端を越えたならば、入力音響信号の全体を探索したことになるので、探索処理を終了する。

【0132】次に、本実施形態による高速信号探索装置の動作実験例を示す。本発明の効果を調べるため、6時間分のテレビ放送の音響信号を入力信号とし、15秒間のCMの音響信号を参照信号として探索実験を行った。本装置は、精度に関しては参照信号を別々に照合した場合と同一なので、照合回数のみを比較する。また、類似値が閾値を上回る場合については必ず全探索を行うので、ここでは、類似値が閾値以下であった照合回数が、参照信号を別々に照合した場合に比べてどのくらいの割合であるか（本出願ではこれを照合回数比という）を調べる。探索のパラメータは、サンプリング周波数=11.025kHz、帯域フィルタの数=7、周波数分析の分析窓長=60ms、分析窓シフト=10ms、各特徴次元におけるビン数=3、時間窓分割数=1、閾値 $\theta = 0.8$ とした。探索結果の一例を図17に示す。縦軸は類似値、横軸は時刻を示す。また、○印は検出箇所を示し、破線は閾値 $\theta$ （0.8）である。

【0133】図18に実験結果を示す。図18の実験例（a）は、無作為に選択した異なる商品のCM3本を参照信号とした場合、実験例（b）は、音が極めて類似している同一商品のCM3本を参照信号とした場合の結果である。参照信号間の類似値は、実験例（a）の場合0.11、0.22、0.23、実験例（b）の場合0.72、0.75、0.88であった。実験例（a）の場合には、照合回数削減の効果は少ないが（照合回数比99.9%）、実験例（b）の場合には、閾値以下の類似値を与えた照合回数が2/3以下となっている（照合回数比62.7%）。このように、本装置は、参照信号どうしの類似値が高い場合に特に有効である。

【0134】§4. 補足

以上、この発明の実施形態を図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の設計の変更等があってもこの発明に含まれる。

【0135】例えば、上記第1～3実施形態では、「ヒストグラムを用いた類似値計算」と「類似値に基づく移動可能量計算」とを特徴とする高速信号探索方法について説明したが、本発明による高速信号探索方法は、必ずしも、この2つの特徴を共に利用する必要は無く、いず

れか一方を利用するだけであっても、従来技術と比べて、格別な効果が得られる。

【0136】また、上記第1～3実施形態では、処理対象信号として音響信号を使用した。図13、14に関する実験例で述べたように、画像の色情報(R、G、B等)を示す信号等を用いてもよい。

【0137】また、特徴量の一例として、第1実施形態では零交差数を用い、第2、3実施形態ではスペクトル特徴を用いたが、本発明で用いられる特徴量はこれらには限定されず、他の特徴量を用いることができる。

【0138】また、上述した各実施形態では、ヒストグラムの類似値に基づいて信号検出を行う場合について述べてきたが、類似値の代わりに距離(ユークリッド距離、L1距離など)を用いて信号検出を行ってもよい。その場合は、類似値に基づく場合における値の大小関係を反転すれば、同等の効果が得られることはいうまでもない。

【0139】また、上述した各実施形態では、高速信号探索装置を単体で示したが、該装置を利用してビデオ自動制御装置を構成することもできる。即ち、本高速信号探索装置をビデオ装置と組み合わせると共に、本高速信号探索装置によって特定の音響信号(テーマソング等)あるいは特定の画像パターン等の発生を検出し、その検出結果に応じてビデオ装置の録画機能等を制御する制御手段を設ける。これによって、特定の番組の開始と終了等にあわせて自動的にビデオ録画機能を起動し、あるいは録画機能を停止し、あるいはインデックス情報を記録したりするビデオ自動制御装置を実現することができる。

【0140】さらに、これによって、放送の音響信号の中から特定のコマーシャルの放映された時刻を検出して自動記録をしたり、特定のテーマソングを検出してビデオ録画を開始したり停止したりすることが可能となる。又、放送から拍手音の発せられた時刻や、笑い声の発せられた時刻などを自動的に監視したり、特定のシーンを検索したりすることも可能となる。更に、これらの処理は音響信号だけでなく、画像パターンなど一般の信号の検出にも応用できる。

【0141】

【発明の効果】以上説明したように、本発明によれば、入力信号に注目窓を設定してその注目窓内の入力信号について特徴量系列を生成し、これとあらかじめ登録された参照信号の特徴量系列との類似値計算を行い、類似値の値に基づいて注目窓の位置の移動量を効率良く変化させながら上記特徴量系列生成と類似値計算の処理を繰り返し、計算した類似値と閾値との比較によって入力信号中の参照信号位置を検出することとしたので、入力信号の全領域に対して注目窓を細かく移動させなくとも、全領域に対して細かく移動させた場合と同等の検出結果を高速で得ることができる。また、本発明によれば、放送

などの長時間にわたる信号の中から特定の音楽やコマーシャルの出現時刻を検索することや、信号データベースなどから特定の信号を含む信号を検索することも可能となる。

【0142】また、本発明によれば、複数の注目窓内の特徴量系列について、これらの特徴量系列間の(時間軸上における)前後関係を考慮に入れて検索を行うことができる。これにより、雑音等によって信号が変動する場合や、平均的な特徴では信号を区別し難い場合等であっても、より高い精度で信号を検出できる。尚、上記「複数の注目窓」の設定形態としては、例えば、もとの注目窓を分割して複数の注目窓として設定すれば、時間の前後関係を考慮に入れて検索を行うことができることになる。

【0143】さらに、本発明によれば、多数の参照信号に基づいて信号を検出する場合であっても、より少ない処理で信号を検出でき、信号検出処理における計算効率の向上が図れる。なお、本発明の記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによっても、より少ない処理で信号を検出して信号検出処理における計算効率の向上を図ることができる。

【図面の簡単な説明】

【図1】 本発明の第1実施形態による高速信号探索装置の構成例を示すブロック図である。

【図2】 本発明の第1実施形態による高速信号探索装置の動作例を示すフローチャートである。

【図3】 特徴抽出装置104の構成例を示すブロック図である。

【図4】 窓の移動前後における類似値計算の一例を示す説明図である。

【図5】 本発明の第1実施形態による高速信号探索装置の検出精度の一例を示す図表である。

【図6】 本発明の第1実施形態による高速信号探索装置の検出速度(総当たり法との速度比)の一例を示す図表である。

【図7】 本発明の第2実施形態による高速信号探索装置の構成例を示すブロック図である。

【図8】 本発明の第2実施形態による高速信号探索装置の処理例を示す説明図である。

【図9】 本発明の第2実施形態による高速信号探索装置の動作例を示すフローチャートである。

【図10】 スペクトル特徴を抽出するための構成例を示すブロック図である。

【図11】 従来方法による音響信号検出結果の一例を示すグラフである。

【図12】 本発明の第2実施形態による音響信号検出結果の一例を示すグラフである。

【図13】 従来方法による画像パターン検出結果の一例を示すグラフである。



【図14】 本発明の第2実施形態による画像パターン検出結果の一例を示すグラフである。

【図15】 本発明の第3実施形態による高速信号探索装置の構成例を示すブロック図である。

【図16】 本発明の第3実施形態による高速信号探索装置の動作例を示すフローチャートである。

【図17】 本発明の第3実施形態による音響信号検出結果の一例を示すグラフである。

【図18】 本発明の第3実施形態における参照信号同士の類似値と照合回数の削減効果との関係例を示すグラフである。

【符号の説明】

- 101 特徴テンプレート作成装置
- 102 音響信号検出装置
- 103 音響信号入力装置
- 104 特徴抽出装置
- 105 ヒストグラム作成装置
- 106 音響信号入力装置
- 107 注目領域走査抽出装置
- 108 特徴抽出装置
- 109 ヒストグラム作成装置
- 110 類似値計算装置
- 111 上限値計算装置
- 112 検出判定装置
- 113 フレーム分割器
- 114 零交差数計数器
- 115 微分器
- 116 零交差数計数器
- 117 微分器

- \* 118 零交差数計数器
- 201 参照特徴量抽出手段
- 202 入力特徴量抽出手段
- 203 類似値計算手段
- 204 移動量計算手段
- 301 信号検出装置
- 310 参照特徴量計算手段
- 312 入力特徴量計算手段
- 314 参照信号間類似値計算手段
- 316 類似値計算手段
- 318 移動可能量計算手段

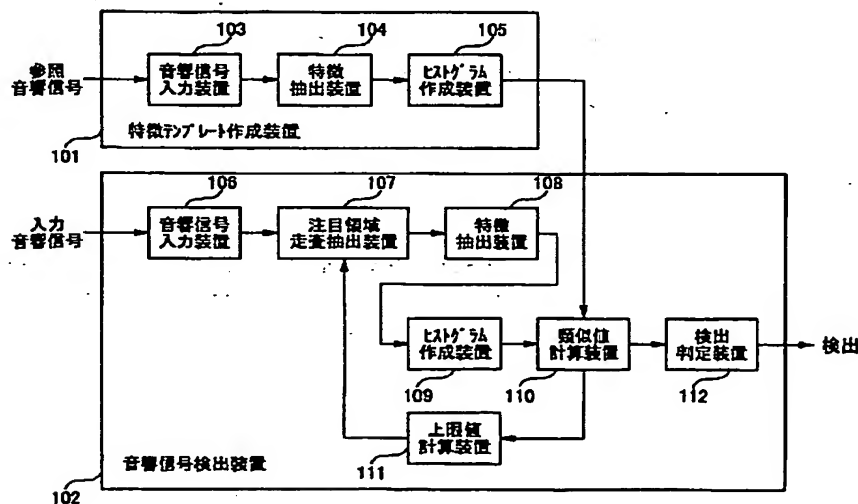
【要約】

【課題】 入力信号の全領域に対して注目窓を細かく移動させなくとも、全領域に対して細かく移動させた場合と同等の検出結果を得ることができる高速信号探索方法、装置およびその記録媒体を提供する。

【解決手段】 特徴テンプレート作成装置101が参照音響信号の特徴量系列を生成し、音響信号検出装置102が入力音響信号に注目窓を設定して参照音響信号に類似する音響信号の存在位置を検出する。この検出処理で音響信号検出装置102は、注目窓内の入力音響信号の特徴量系列を生成する過程と、参照音響信号の特徴量系列と生成した入力音響信号の特徴量系列との類似値を計算する過程と、該類似値に基づいて注目窓の移動可能量を計算する過程と、該移動可能量に基づいて注目窓を設定する過程とを繰り返す。これにより、注目窓の各位置について類似値を計算し、閾値との比較によって参照信号の存在を判定する。

\*

【図1】



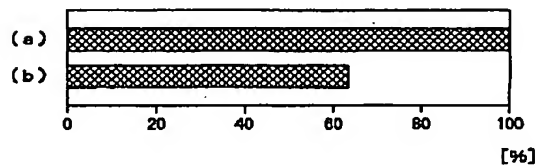
【図5】

窓長	再現率	精度
5.94 秒	0.99	0.97
11.89 秒	1.00	1.00

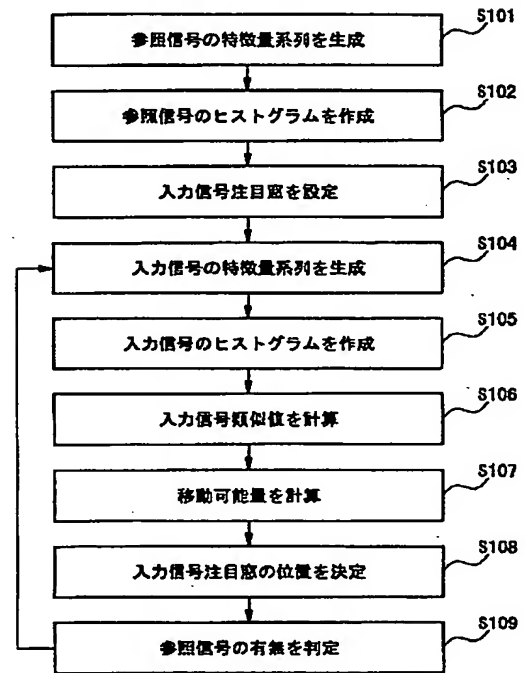
【図6】

窓長	総当たり法との速度比
5.94 秒	22.8
11.89 秒	19.1

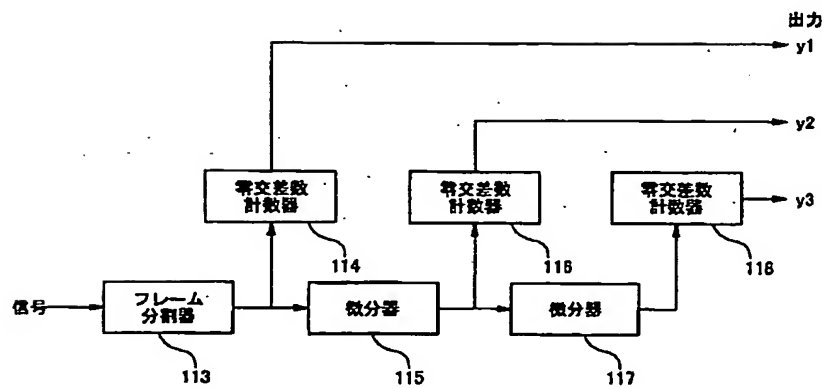
【図18】



【図2】

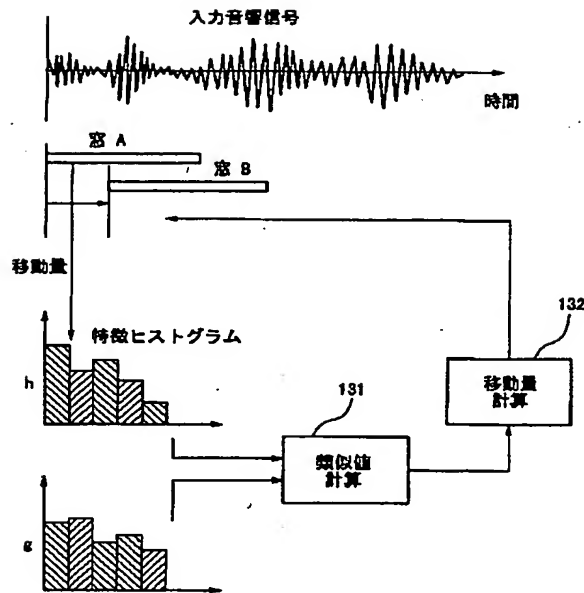


【図3】

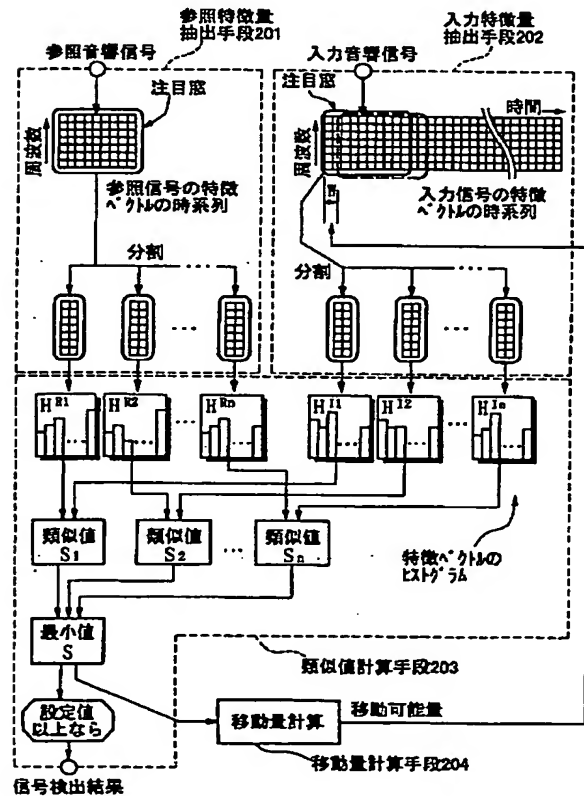


104:特徴抽出装置

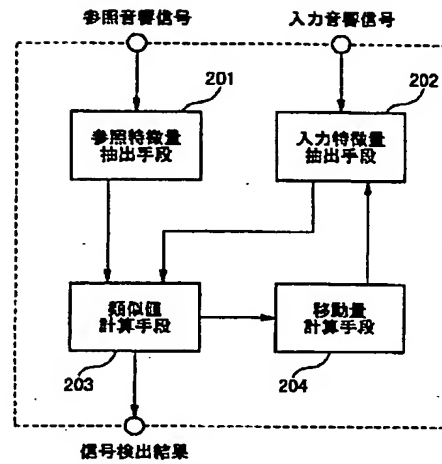
【図4】



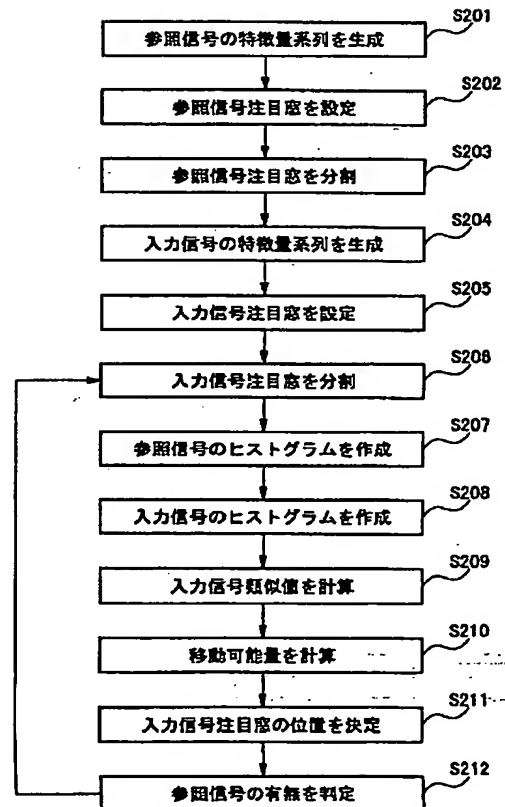
【図8】



【図7】



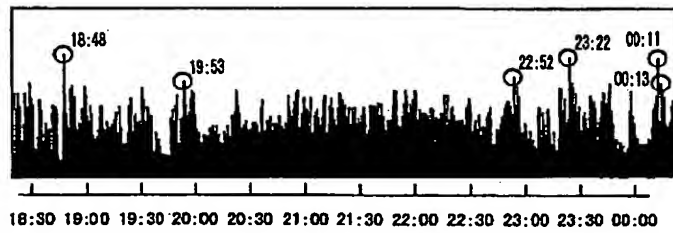
【図9】



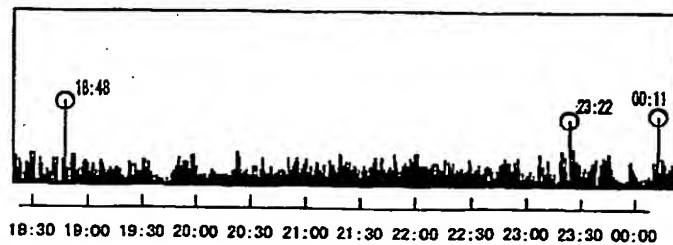
【図10】



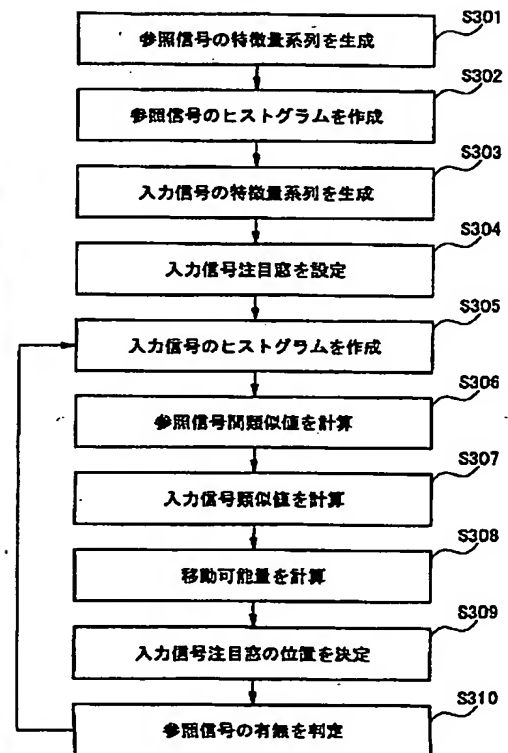
【図11】



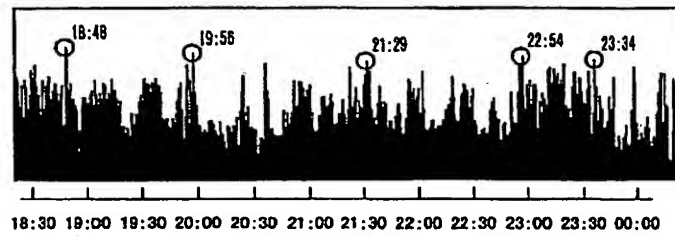
【図12】



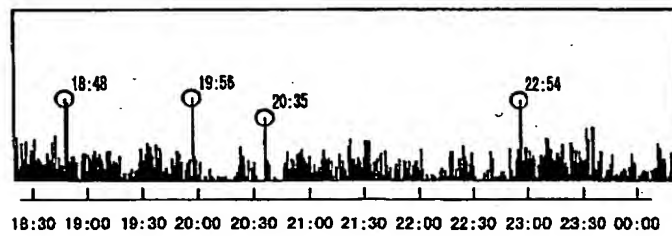
【図16】



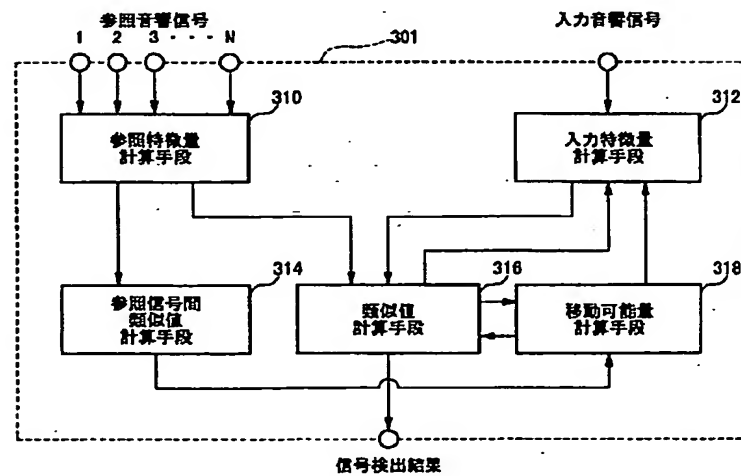
【図13】



【図14】

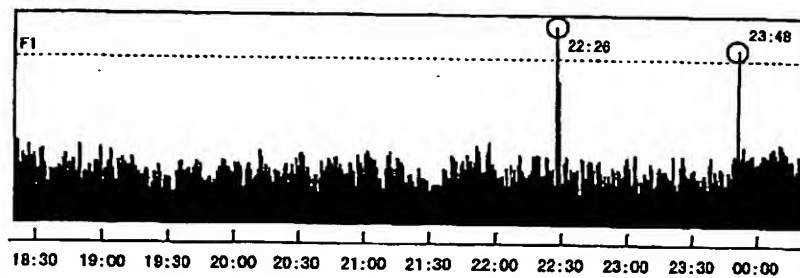


【図15】





【図17】



フロントページの続き

特許法第30条第1項適用申請有り Proceedings of IEEE 1998 International Conference on Acoustics, Speech and Signal Processing, Vol. 6, "Quick Audio Retrieval Using Active search" p. 3777-3780, May 12-15, 1998, Seattle, Washington, USAに発表

- (56)参考文献 特開 平4-198999 (J P, A)  
 特開 平10-136297 (J P, A)  
 電子情報通信学会技術研究報告 [パターン認識・メディア理解] Vol. 97, No. 40, PRMU97-17, 「局所色情報を用いた高速物体探索-アクティブ探索法」 p. 127-134 (1997年5月15日発行)  
 電子情報通信学会技術研究報告 [パターン認識・メディア理解] Vol. 98, No. 274, PRMU98-80, 「マルチモーダルアクティブ探索を用いた画像・音響時系列の高速探索」 p. 51-58 (1998年9月17日発行)  
 電子情報通信学会技術研究報告 [音声] Vol. 99, No. 120, SP99-23, 「時系列アクティブ探索法における音響信号の高速AND/OR探索」 p. 41-48 (1999年6月17日発行)  
 電子情報通信学会技術研究報告 [音声] Vol. 98, No. 611, SP98-141, 「セグメントの高速探索法」 p. 41-48 (1999年2月19日発行)  
 電子情報通信学会論文誌 Vol. J81-D-▲II▼ No. 9, September 1999, 「局所色情報を用いた高速物体探索-アクティブ探索法-」, p. 2035-2042, (平成10年9月25日発行)  
 日本音響学会平成10年度秋季研究発表会講演論文集 2-9-24「ヒストグラム特徴系列に基づく長時間音響信号の高速探索」 p. 561-562 (平成10年9月24日発行)  
 日本音響学会1999年度春季研究発表会講演論文集 1-6-9「複数の音響信号の高速探索-参照信号のOR探索に於ける照合回数の削減-」 p. 447-448 (平成11年3月10日発行)

(58)調査した分野(Int.Cl.<sup>7</sup>, DB名)

G10L 15/08  
 G10L 15/00  
 G10L 15/10  
 H04N 5/928  
 INSPEC (DIALOG)  
 JICSTファイル (JOIS)  
 WPI (DIALOG)